

Rich-cores in networks

Athen Ma and Raúl J Mondragón
Queen Mary University of London,
School of Electronic Engineering and Computer Science,
London, United Kingdom

February 19, 2014

Abstract

A core is said to be a group of central and densely connected nodes which governs the overall behavior of a network. Profiling this meso-scale structure currently relies on a limited number of methods which are often complex, and have scalability issues when dealing with very large networks. As a result, we are yet to fully understand its impact on network properties and dynamics. Here we introduce a simple method to profile this structure by combining the concepts of core/periphery and rich-club. The key challenge in addressing such association of the two concepts is to establish a way to define the membership of the core. The notion of a “rich-club” describes nodes which are essentially the hub of a network, as they play a dominating role in structural and functional properties. Interestingly, the definition of a rich-club naturally emphasizes high degree nodes and divides a network into two subgroups. Our approach theoretically couples the underlying principle of a rich-club with the escape time of a random walker, and a rich-core is defined by examining changes in the associated persistence probability. The method is fast and scalable to large networks. In particular, we successfully show that the evolution of the core in *C. elegans* and World Trade networks correspond to key development stages and responses to historical events respectively.

Many networks exhibit a core/periphery dichotomy which is important for the understanding of network properties and dynamics (for a review see ref. [1]). For instance, this meso-scale structure helps explain elitism in economic growth among countries [2], organization of the human brain dynamics [3] and inter-hub configurations in communication networks [4]. The constitution of a core/periphery structure often refers to a simple two-class partition [5, 6, 7]. Members of the core are high degree nodes which have high inter-core and intra-partition linkages. In the original definition, the core/periphery was obtained by applying block-modeling to the adjacency matrix, where the connectivity within the periphery was assumed to be non-existent [5]; however, this assumption is too stringent for most real networks, and newer methods tend to relax this constraint. Other methods include using closeness centrality to minimize the diameter between the core and the rest of the network [6, 7], and applying Markov chains to describe random walks and provide an implicit partition [8]. Most studies focus on a single core, while the possibility of having a multi-core has been discussed [5, 9], the number of available methods for their profiling is rather limited.

The notion of a rich-club is used to describe the connectivity between high degree nodes, and it has been applied to profile meso-scale properties in networks by examining the density of connections between high degree nodes [10, 11, 12, 13, 14, 15]. A rich-club influences the functionality of a network, as demonstrated in the transmission of rumors in social networks

[16], the delivery of information in the Internet [10] and its strong influence on both the network assortativity and transitivity [12]. Interestingly, the presence of a rich-club naturally divides a network into two parts. This means that the rich-club concept coincides with two fundamental prerequisites of a core/periphery structure: *high degree nodes* and *two-class partition*. However, there is, at present, no general method to define the membership of the club [12, 14, 17].

Here, we present a *rich-core* method which detects the core/periphery structure in complex networks, and formally defines the members of the associated rich-club. The method marks the formation of a rich-core by examining the persistence probability of random walks among high degree nodes. We analyzed a wide range of networks and successfully identified the core among them; our method also reveals multi-core within a core structure where one is present. Networks with a temporal nature have also been studied and our results show that the evolution of a core closely reflects the development and re-alignment of relationships over time in real networks. For example, changes in the core coincide with key development stages in a biological process and with timing of major historical events in trade development. In addition, anomalous nodes in networks can be uncovered with reference to null models by discriminating statistical differences. The rich-core method is simple, fast and applicable to very large networks.

1 Results and Discussion

Consider an unweighted and undirected graph. We rank the order of importance of nodes in descending order of their degree, such that the node with the highest degree is ranked first and so on. For a given node, we divide its links into two groups: those with nodes of a higher rank and those with a lower rank. More formally, a node with a rank r has degree k_r ; the number of links it shares with nodes of a higher rank is k_r^+ and the number of links with nodes of a lower rank is $k_r - k_r^+$. Core nodes are often high degree nodes that are densely connected with each other [5], and we assume that the connectivity of a highly ranked node with other *higher* ranked nodes contributes towards the constitution of a core. Similarly, if such a node has very few links with higher ranked nodes, it is likely to be a member of the periphery. To detect the core we propose this straightforward procedure (see Materials and Methods for full technical description). Starting from the node with the highest rank, as r increases the number of links k_r^+ that node r shares with nodes of a higher rank fluctuates. There will be a node r^* where k_r^+ has reached its maximum, and from that node onwards k_r^+ is always less than $k_{r^*}^+$. This change in the connectivity among the highly ranked nodes is defined as the boundary of the (rich) core; the nodes with a rank less than or equal to r^* are the core and the rest belong to the periphery. We studied the Zachary Karate Club network [18] and the rich-core is formed by the 10 highest ranked nodes (Fig. 1).

This pragmatic way of defining a core is related to the concept of a random walker in a network. Consider a network which is partitioned into two sets: S_c is the core and S_p is the periphery. A random walker jumps from one node to another following a link between any pair of nodes, and the probability for the walker to visit a given node is proportional to the node's degree. The time it takes a random walker to escape from S_c to S_p is τ_c . If S_c is the set of nodes with the highest rank, as we include another node into this set the escaping time will naturally increase. Eventually, S_c will contain all the nodes in the network, and the escaping time will diverge as the random walker is always contained in this set and has nowhere to escape to. This means that if we begin by putting only the top ranked node in set S_c , and gradually increase the size of S_c by adding nodes in decreasing order of their rank, the escaping time will always increase. If we consider the rate of change of the escaping time as nodes are being added to the set, the boundary of a rich-core depicts the point in which the rate of increase on the escaping

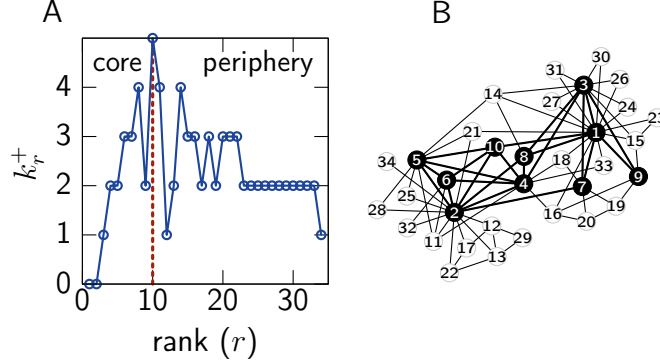


Figure 1: The rich-core for the Zachary Karate Club network. (A) The number of links k_r^+ that node r shares with nodes of a higher rank. The boundary of the core is marked by $k_{r^*}^+$ where k_r^+ is at its maximum, as indicated by the dotted line. (B) A graphical representation of the network with core (black) and periphery (white) nodes derived from the rich-core method.

time changes from a fast to a slow pace (Materials and Methods).

This definition of a rich-core can be easily extended to weighted networks [19, 20]. Consider w_{min} is the minimal weight linking two nodes in a network and the link between nodes i and j has a weight of w_{ij} . This link is represented by $\lceil w_{ij}/w_{min} \rceil$ links and the ranking is performed in units of the minimal weight. Each node is assigned to $\sigma_i = \lceil \sum_j w_{ij}/w_{min} \rceil$ links, and part of this quantity arisen from the node's linkage to nodes of a higher rank is referred to as σ_r^+ ; similarly, the remaining proportion, $\sigma_r - \sigma_r^+$, is the normalized weight that node r shares with nodes of a lower rank. The core boundary is node r^* such that $\sigma_{r^*}^+ > \sigma_{r^*+i}^+$ for $i = 1, \dots, N - r$.

Similarly, the notion of a rich-core can equally be applied to directed networks by dividing links into in- and out- links and quantifying their corresponding weights. In this case the definition of a rich-core does not only depend on the weight of the nodes but also the direction of their links (in- or out-degrees). An example is the assessment of web pages using PageRank where it has been observed that the popularity of a node is closely related to its in-degree [21]. Let $\sigma_r^{+in} + \sigma_r^{+out}$ be the total strength of interactions between node r and the nodes $r' < r$, the core boundary is the node r^* such that $\sigma_{r^*}^{+in} + \sigma_{r^*}^{+out} > \sigma_{r^*+i}^{+in} + \sigma_{r^*+i}^{+out}$ for $i = 1, \dots, N - r$.

We analyzed the World Trade network [22] in which nodes are countries and links are trade channels; the latter can represent the direction of trade to specify an import or export relationship. The associated financial value can be seen as the weight of a given link. The overall connectivity of the network is found to be high as countries are interrelated in many ways. We first examined the network as a binary and unweighted network and a link simply refers to the presence of trade. A total of 60% of all the countries in the world are part of this network as a result of the globalization of trading at the time [23]. By ranking countries in descending order of their degree, Fig. 2A shows the number of trade relationships each of these countries has with countries of a higher rank in 1990 and the core consists of 106 countries. This interpretation of the network confirms close trading relationships found among the countries within the European Union (EU), which are highly ranked. The USA, which is a dominant importer and exporter by value, is ranked 15. Now, if we take both the direction and weight into consideration and nodes are ranked in the descending order of their exports (Fig. 2B), members found in the core are consistent with the top importers and exporters in the world at the time [23].

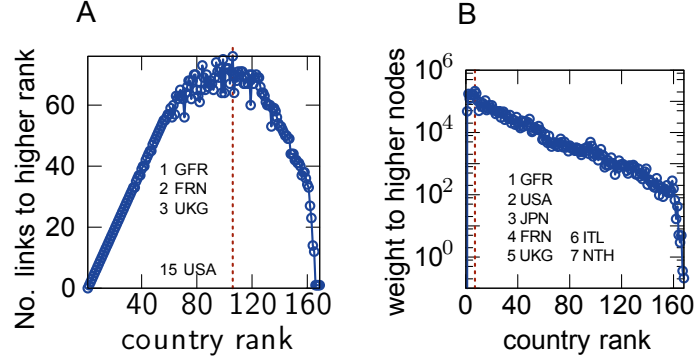


Figure 2: The World Trade network in year 1990. (A) The unweighted undirected network which represents trade relationships between the countries. The core has 106 countries. (B) The weighted directed network representing the exchange of wealth between the countries. The core has seven countries.

1.1 Evolution

Networks are often found to be temporal in nature as they are subject to formation, dissolution and rewiring of links [24, 25]. Continuing with the example of World Trade we studied the core in the period between 1948 to 2000. Applying our method results in a core which consists of very selective group of countries, corresponds to 4% to 6% of all the countries (Fig. 3A). This can be explained by referring to the way in which the World Trade has grown since the Second World War. International trading is said to be growing steadily but unevenly since the 1940s, as trade barriers were imposed by events such as the Cold War. The network is also strongly influenced by other key historical events, geographical distance, composition (e.g. products and services) and nature of trade [23]. Throughout the 1980s and 1990s, we see substantial reduction in the cost of shipping due to the explosion of air freight, the collapse of the Soviet Union has led to many independent countries, and industrialization of developing countries; all these events have shaped the development of trade worldwide, leading to a great leap in globalization. This is somewhat reflected by the way the membership of the core changes over time. The USA, Germany, Japan, France and the United Kingdom have been the top importers and exporters in the world during the period of the study and it can be seen that these countries have been members of the core for the entire time. Canada has been drifting in and out of the top ten in the World Trade ranking during the same period, and we can see a similar variability in its core membership. In addition, the economic reform in China, which started in the late 1970's, has led to a steady growth of $\sim 9\%$ in World Trade per year, and our results illustrate that China became a member of the core in 1997. In fact, China joined the World Trade Organization in 2001.

Another example of how changes in a core tie in with key events in real networks can be found in the development of *C. elegans* neuronal connections (Fig. 3B). Almost the entire core is developed within the first 500 minutes [14]. The formation of the core coincides with Embryogenesis, and it has been suggested that the highly connected neurons appear in the early development to minimize the energy cost by creating the core connections among key nodes that are not physically far apart; these connections can then be extended during the process of body elongation. New neurons are found after hatching in the late L1 larval stage at approximately

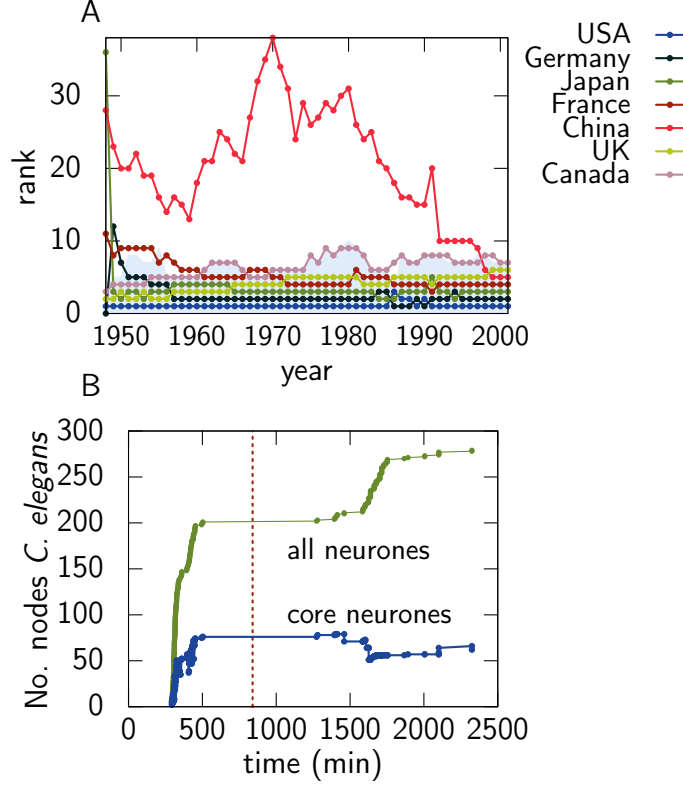


Figure 3: (A) Evolution of the core for the World Trade network. The shaded area (light blue) is the size of the core. The countries shown are the members of the core in year 2000. (B) Evolution of the core for *C. elegans*. The red dotted line marks the hatching time.

1250 minutes and the total number of neurons continues to grow until the start of L4 larval stage at approximately 2400 minutes [26]. The post-hatching development causes the relative connectivity among the existing core neurons to decrease, resulting in a reduction in the overall size of the core. The shrinking of the core coincides with the timing of Gonadogenesis.

1.2 Multi-core and Relative Size

A network can have a multi-core which consists of subgroups of nodes or disjoint parts. The former refers to a core which has a well-defined internal structure such that cliques can be observed within its configuration. A simple way to divide them into their prospective groups is to re-use the ranking of the nodes (Materials and Methods). We analyzed the network of scientists [27] (NetSci) by firstly defining the core of the network (Fig. 4). Secondly, we examined the internal structure of the core by employing the algorithm for subgroup classification, and our results reveal that the core can be divided into well defined subgroups. On the other hand, the latter can be assessed using reachability and can be achieved using a shortest-path algorithm.

The size of a core is an important property of a network as it has been suggested that a sizable core makes a network more flexible and adaptable to changes [28], and a small core makes a network more controllable [29]. The relative size of a core $c = N_c/N$ is the ratio between

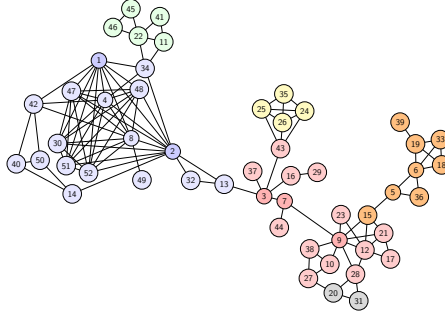


Figure 4: Core of the NetSci network and its internal structure. The multi-core found is further subdivided into their corresponding parts and highlighted in different colors.

the number of nodes in the core N_c relative to the total number of nodes in the network N . A network that has no periphery will have a relative core of 1, e.g. a fully connected network; a star network with N nodes has a relative core of $1/N$ and a core-less network has no links. We examined the relative size of a range of complex networks, ranging from man-made, social and biological, and the cores observed in these networks have different size and structure, in part reflecting their functionality. For instance, the Amazon and the Internet networks both have a relatively small core. The former is found to be disjoint, and as the network contains information about products recommendations, the results provide evidence of efficient information transfer within the network but only restricted to localized parts. The latter has a well connected single core which reflects its design for efficient routing between Autonomous System domains. The *C. elegans* neuronal network [34] has a relative large single core, and this perhaps reflects the adaptability of the neuronal network to living. The CA-road network [32] has a relatively large multi-core which represents the existence of many crossroads, providing great flexibility in route choices as they present many possibilities between different geographical points. We did not observe any characteristic size of the core related to the kind of networks, i.e. man-made, social or biological.

1.3 Anomalies in the core and its nodes

Our rich-core method can be used to detect anomalies with respect to the core size and connectivity among members. We employed a randomization method to create an ensemble of networks which in turn are used as a reference null model for comparison purposes. The randomization is restricted to preserve the ranking of the nodes, that is the weight/degree of the nodes (Materials and Methods). Using *C. elegans* as an example, we created an ensemble of 100 networks with the same degree distribution as the original network. From the ensemble we evaluated the average number of links that node r has with nodes of a higher rank, i.e. $\langle k_r^+ \rangle$, and the standard deviation of this quantity. Fig. 6A shows the number of links between a node of rank r and a node of rank $r' < r$ and the average number of links obtained from the null-model. The boundary of the core-periphery is the dotted blue line. The red line is $\langle k_r^+ \rangle$ and the pink shaded area demarcates two standard deviations from the mean value. This implies that nodes outside the shaded area can be considered anomalous. For example, node 14 which corresponds to neuron RIAR, has an anomalous connectivity as it only shares one of its links with nodes of a higher rank. The randomization can also be used to decide if the size of a core is within the expected value. Fig. 6B shows the average size of the core obtained from the ensemble of networks (red dotted line) and

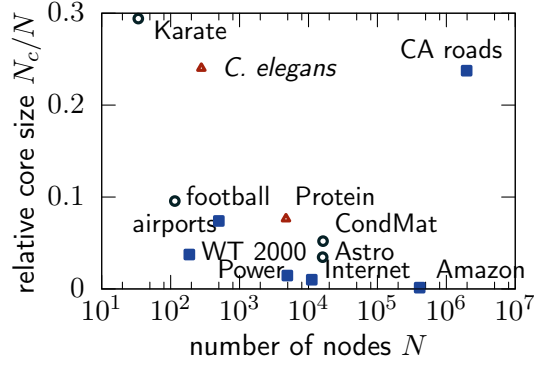


Figure 5: Relative size of the core in different kinds of real networks. Man-made (square): US airports (airports) [30], Amazon recommendation (Amazon) [31], Californian (CA) roads [32], Internet [33], Power grid (Power) [34] and World Trade (WT) 2000 [22]. Social (circle): Astrophysics collaborations (Astro) [35], Condensed Matter collaborations (CondMat) [35], American College football (football) [36] and the Zachary Karate club (karate) [18]. Biological (triangle): *C. elegans* [34] and Protein [37].

the pink area corresponds to two standard deviations from this mean. The dotted blue line is the size of the core obtained from the empirical data; as this value falls inside the pink area we can conclude that the size of the core for the *C. elegans* is what we expected.

2 Conclusions

We developed a method to profile core/periphery in networks by theoretically coupling the notion of a rich-club and the persistence probability of random walks, resulting in finding a rich-core. The method is simple, fast and applicable to large networks, as demonstrated by the range of artificial, social and biological networks. Our theoretical work shows that the method is robust and can be extended to both directed and weighted graphs. Furthermore, we uncovered important characteristics in real networks through the profiling of the core/periphery structure, contributing towards the understanding of the significance of such meso-scale property. In particular, development and realignment of relationships are often found in networks with temporal nature, and our studies on both the World Trade and *C. elegans* networks showed that the evolution of the core closely tie in with the timing of major historic events and key development stages respectively. We also demonstrated how the method reveals the organization of the core in a microscopic scale and characterizes multi-core. We showed an effective way to identify any anomalies in the core by comparing the membership of the core of a given network with its null counterparts, and how this approach can be used to evaluate the expected size of of a given core.

We thank J. Schormans for his comments on the manuscript, X. Lu and C. Gray for early discussions related to this research.

References

- [1] Csermely P, London A, Wu LY, Uzzi B (2013) Structure and dynamics of core-periphery networks. *arXiv preprint arXiv:1309.6928*.

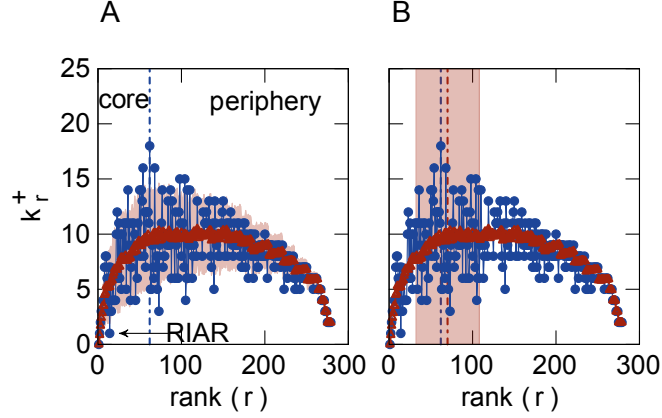


Figure 6: (A) Comparison on the connectivity of the core between the empirical data and the null model for *C. elegans*. The shaded area marks two standard deviations from the sample mean. The neuron RIAR is found to have an anomalous connectivity with the highest ranked nodes. (B) Comparison on the core size between the empirical data (blue line) and the null model (red line). Similarly, the shaded area shows two standard deviations.

- [2] Krugman P, Venables AJ (1995) Globalization and the inequality of nations. *The Quarterly Journal of Economics* 110:857–880.
- [3] Bassett DS, et al. (2012) Core-periphery organization of human brain dynamics. *arXiv preprint arXiv:1210.3555*.
- [4] Carmi S, Havlin S, Kirkpatrick S, Shavitt Y, Shir E (2007) A model of internet topology using k-shell decomposition. *Proceedings of the National Academy of Sciences* 104:11150–11154.
- [5] Borgatti SP, Everett MG (1999) Models of core/periphery structures. *Social networks* 21:375–395.
- [6] Holme P (2005) Core-periphery organization of complex networks. *Physical Review E* 72:046111.
- [7] Da Silva MR, Ma H, Zeng AP (2008) Centrality, network capacity, and modularity as parameters to analyze the core-periphery structure in metabolic networks. *Proceedings of the IEEE* 96:1411–1420.
- [8] Della Rossa F, Dercole F, Piccardi C (2013) Profiling core-periphery network structure by random walkers. *Scientific reports* 3.
- [9] Puck Rombach M, Porter MA, Fowler JH, Mucha PJ (2012) Core-periphery structure in networks. *arXiv preprint arXiv:1202.2684*.
- [10] Zhou S, Mondragón R (2004) Accurately modeling the internet topology. *Physical Review E* 70:066108.
- [11] Colizza V, Flammini A, Serrano MA, Vespignani A (2006) Detecting rich-club ordering in complex networks. *Nature* 2:110–115.

- [12] Xu XK, Zhang J, Small M (2010) Rich-club connectivity dominates assortativity and transitivity of complex networks. *Physical Review E* 82:046117.
- [13] van den Heuvel MP, Sporns O (2011) Rich-club organization of the human connectome. *The Journal of Neuroscience* 31:15775–15786.
- [14] Towilson EK, Vértés PE, Ahnert SE, Schafer WR, Bullmore ET (2013) The rich club of the *C. elegans* neuronal connectome. *The Journal of Neuroscience* 33:6380–6387.
- [15] Mondragón RJ (2012) Network null model based on maximal entropy and the rich-club. *arXiv preprint arXiv:1210.2211*.
- [16] Masuda N, Konno N (2006) Vip-club phenomenon: Emergence of elites and masterminds in social networks. *Social Networks* 28:297–309.
- [17] Valverde S, Solé R (2007) Self-organization versus hierarchy in open-source social networks. *Physical Review E* 76:046118.
- [18] Zachary WW (1977) An information flow model for conflict and fission in small groups. *Journal of anthropological research* pp 452–473.
- [19] Serrano MA (2008) Rich-club vs rich-multipolarization phenomena in weighted networks. *Physical Review E* 78:026101.
- [20] Zlatic V, et al. (2009) On the rich-club effect in dense and weighted networks. *The European Physical Journal B* 67:271–275.
- [21] Fortunato S, Boguñá M, Flammini A, Menczer F (2008) Approximating PageRank from in-degree. *Lecture Notes in Computer Science* 4936:59.
- [22] Gleditsch KS (2002) Expanded trade and GDP data. *Journal of Conflict Resolution* 46:712–724.
- [23] Report WT (2013) World Trade Report 2013. *World Trade Organization*
- [24] Leskovec J (2008) *Dynamics of large networks* (ProQuest).
- [25] Kostakos V (2009) Temporal graphs. *Physica A: Statistical Mechanics and its Applications* 388:1007–1023.
- [26] Byerly L, Cassada R, Russell R (1976) The life cycle of the nematode *caenorhabditis elegans*. i. wild-type growth and reproduction. *Developmental Biology* 51:23.
- [27] Newman ME (2006) Finding community structure in networks using the eigenvectors of matrices. *Physical Review E* 74:036104.
- [28] Csete M, Doyle J (2004) Bow ties, metabolism and disease. *TRENDS in Biotechnology* 22:446–450.
- [29] Liu YY, Slotine JJ, Barabási AL (2011) Controllability of complex networks. *Nature* 473:167–173.
- [30] Colizza V, Pastor-Satorras R, Vespignani A (2007) Reaction–diffusion processes and metapopulation models in heterogeneous networks. *Nature Physics* 3:276–282.

- [31] Leskovec J, Adamic LA, Huberman BA (2007) The dynamics of viral marketing. *ACM Transactions on the Web (TWEB)* 1:5.
- [32] Leskovec J, Lang KJ, Dasgupta A, Mahoney MW (2009) Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Mathematics* 6:29–123.
- [33] Chen Q, Chang H, Govindan R, Jamin S (2002) *The origin of power laws in Internet topologies revisited* (IEEE), Vol. 2, pp 608–617.
- [34] Watts DJ, Strogatz SH (1998) Collective dynamics of small-world networks. *Nature* 393:440–442.
- [35] Newman ME (2001) The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences* 98:404–409.
- [36] Girvan M, Newman ME (2002) Community structure in social and biological networks. *Proceedings of the National Academy of Sciences* 99:7821–7826.
- [37] Colizza V, Flammini A, Maritan A, Vespignani A (2005) Characterization and modeling of protein–protein interaction networks. *Physica A: Statistical Mechanics and its Applications* 352:1–27.
- [38] Piccardi C (2011) Finding and testing network communities by lumped markov chains. *PloS One* 6:e27028.
- [39] Maslov S, Sneppen K, Zaliznyak A (2004) Detection of topological patterns in complex networks: correlation profile of the internet. *Physica A: Statistical Mechanics and its Applications* 333:529–540.

Methods

Rich-club Coefficient and Core/Periphery Profile

To find the boundary of a rich-core, we examine the escaping time it takes a random walker to leave a core. The escaping time is related to the notion of persistence probability α which indicates the cohesiveness in a subgraph [8, 38]. The persistence probability of cluster S_C is

$$\alpha_c = \frac{\sum_{i,j \in S_c} \pi_i m_{ij}}{\sum_{i \in S_c} \pi_i}. \quad (1)$$

where π_i is the probability that a random walker is in node i , and m_{ij} is the probability that a random walker moves from node i to node j . The escaping time is $\tau_c = (1 - \alpha_c)^{-1}$. Assuming that α_x can be approximated by a continuous function $\alpha(x) = g(x)/f(x)$ where x is a continuous rank. The first derivative $\alpha'(x)$ increases with x , as the number of nodes rises and eventually diverges to 1 when all the nodes are included. The rate of increase in the persistence probability is given by the second derivative $\alpha''(x) = g''(x)/f(x) + 2g(x)f'(x)^2/f(x)^3 - 2f'(x)g'(x)/f(x)^2 - g(x)f''(x)/f(x)^2$. To first approximation $\alpha''(x) \simeq 0$ if $g''(x) = 0$ as $f(x)$ is a positive increasing function of x .

The rich-club coefficient is defined as [10]:

$$\phi(r) = \frac{2E(r)}{r(r-1)} = \frac{2 \sum_{i=1}^c k_i^+}{r(r-1)} \quad (2)$$

where $E(r)$ is the number of links between the r nodes. For undirected networks α_c is given by the sum of the links between the nodes in S_c divided by the sum of the degrees of the nodes in S_c . If a_{ij} are the elements of the adjacency matrix, then

$$\alpha_c = \left(\sum_{i,j \in S_c} a_{i,j} \right) / \sum_{i \in S_c} k_i = \frac{\sum_{i=1}^c k_i^+}{\sum_{i=1}^c k_i}. \quad (3)$$

Again, assuming that α_c can be approximated with a continuous function $\alpha_c = g(c)/f(c)$ where $g(x) = \int k^+(x)dx$ then $g''(x^*) = 0$ means that $k^+(x^*)$ has a maximum or a minimum at the value x^* , in this case we are interested in the maximum. We refer the point x^* where $g''(x^*) = 0$ as the boundary of the rich-core and nodes in S_x are the members of the core. The explicit relationship between the escaping time and the rich-club coefficient is obtained by substituting Eq. 2 into Eq. 3, giving

$$\alpha_r = \frac{r(r-1)}{2} \frac{\phi(r)}{\sum_{i=1}^r k_i}. \quad (4)$$

For weighted networks, as we are considering the weights as undirected multilinks the same argument applies when defining the core. For directed networks, the persistence probability α_c is given by the ratio between the number of times a random walker transits inside the core and the number of times it visits the core. The former is proportional to the number of links inside the core, which we denoted as $\sigma_{r^*}^{+in} + \sigma_{r^*}^{+out}$. The latter is given by the sum $\sum_{i \in S_c} \pi_i$. The in-degree k_i of node i is assumed to be a good approximation of π_i [21]; hence, the number of times a random walker visits the core increases with r , as the nodes are ranked in decreasing order of their in-degree.

Rich-Core Algorithm

To find the core of a given network,

1. rank nodes in decreasing order of their weight (specifically, degree, in-degree and weight for undirected, directed weighted networks respectively.)
2. evaluate the number of links k^+ between node with rank r and nodes with rank $r' < r$
3. find the boundary of the core, defined by the node where $k_{r^*}^+ > k_r^+$ for $r > r^*$

Rank degeneracy

As numerous nodes can have the same degree, the ranking of nodes with equal degree is not well defined. This degeneracy in the ranking scheme would affect the determination of the boundary nodes, and hence the size of a core. To evaluate the effect of the degeneracy in the definition of a core we randomly re-ranked the nodes with equal degree and measured the change of the core nodes. We observed that this re-ranking only has a minor effect when defining a core.

Multi-core

In undirected networks the ranking of the nodes inside the rich-core can be used to divide the core nodes into subgroups. The algorithm is based on the idea that a random walker that is in a core node of low rank will “drift” to a node of a higher rank, the algorithm evaluates this process in reverse.

1. Starting from the highest ranking node
2. If the node is not a member of a subgroup, create a new subgroup
3. All the neighbors of the selected node which have a lower rank than the selected node belong to the same subgroup.
4. Repeat this until there is no neighbors with a lower rank left. Mark these nodes as members of the same subgroup.
5. If all the nodes of the rich-core have been marked then finish, if not
6. Select the next ranking node, go to step 2

Inside a core there can be nodes which are at the boundary of two subgroups. To decide which subgroup these nodes belong to, we assign the boundary nodes to the subgroup in which they share the highest number of links. If the nodes share the same number of links with both subgroups, the subgroup with the highest ranked nodes will be chosen.

Construction of Null Models

The null model is generated using Zlatic et. al approach [20] which is a generalization of Maslov, Sneppen and Zaliznyak method (MSZ) [39] to generate null models. Zlatic et al. redistributes the weights of the links by preserving the strength of the nodes as follows. If w_m is the minimum value of the weights in the original network then the rewiring is done by changing the weights of two pairs of links by an amount w_m . The rewiring consists of selecting two links at random and exchanging one of the end nodes of the first link with an end node of the second link.